

# Information Flow in Logical Environments

Robert E. Kent

Ontologos

**Abstract.** This paper describes information flow within logical environments. The theory of information flow, the logic of distributed systems, was first defined by Barwise and Seligman [1]. Logical environments are a semantic-oriented version of institutions. The theory of institutions, which was initiated by Goguen and Burstall [5], is abstract model theory. Information flow is the flow of information in channels over distributed systems. The semantic integration of distributed systems, be they ontologies, databases or other information resources, can be defined in terms of the channel theory of information flow. As originally defined, the theory of information flow uses only a specific logical environment in order to discuss information flow. This paper shows how information flow can be defined in an arbitrary logical environment.

**Keywords:** information flow, logical environment, distributed systems, channel, core, information integration, structure, theory and logic

## 1 Introduction

We describe information flow in logical environments. The theory of Information Flow (IF) is [1] the logic of distributed systems. The theory of Institutions (INS) is [5] abstract model theory. Both can be regarded as ways to describe, classify, model, extract, and apply patterns of knowledge in the design and maintenance of ontologies [6], databases [3] and other knowledge resources. The theory IF and the theory INS are related in two ways. On the one hand, IF is a special case of first order logic (FOL) (and FOL is just one institution): the types of IF can be regarded as unary predicates and (then) the instances of IF are to be regarded as individuals that may satisfy these unary predicates. So IF is both a special case and a specific instance of INS. This point of view, advocated in a recent paper by Goguen [4], has been used for several years in the Information Flow Framework (IFF), as evidenced by the various documents and ontologies listed at the IFF website [8]. On the other hand, the IF classification relation (between instances and types) can be regarded as an abstraction of the INS satisfaction relation (between structures and sentences). As such, anything in the theory of IF that is treated abstractly can be used to extend the theory of INS. So we can view IF as an abstraction and extension of INS. This paper argues for this point of view. We first discuss a more semantic form of institutions called logical environments. Then we demonstrate how the basics of information flow can be define in logical environments. We follow the example set by the theory of IF [1] in the use of principles to codify and structure the discussion.

## 2 Logical Environments

### 2.1 Structures

**World Principle:** *The world exists a priori and is describable.*

We represent the world as a category (mathematical context) **Struc** of semantical structures and structure morphisms. We assume the category of structures is self-referential, in that the structures and their morphisms contain (are indexed by) linguistic mechanisms for self-description. This index is represented by a functor (passage between contexts)  $\mathbf{lang} : \mathbf{Struc} \rightarrow \mathbf{Lang}$  from the category of semantical structures to a category **Lang** of (logical) languages and language morphisms. For any structure  $M$ ,  $\mathbf{lang}(M)$  is a language capable of describing  $M$ , and for any structure morphism  $f : M_1 \rightarrow M_2$ ,  $\mathbf{lang}(f) : \mathbf{lang}(M_1) \rightarrow \mathbf{lang}(M_2)$  is a language morphism capable of describing  $f$ . Hence, a world consists of a triple  $\mathbf{Wrld} = \langle \mathbf{Struc}, \mathbf{Lang}, \mathbf{lang} \rangle$  as discussed above.

The institutional approach assumes that multiple worlds exist and are relational. In general, different worlds use different languages for description. Worlds are related through world morphisms. A world morphism  $\mathbf{wrld} = \langle \mathbf{struc}, \mathbf{lang} \rangle : \mathbf{Wrld}_1 \rightarrow \mathbf{Wrld}_2$  consists of a structure functor  $\mathbf{struc} : \mathbf{Struc}_1 \rightarrow \mathbf{Struc}_2$  and a language functor  $\mathbf{lang} : \mathbf{Lang}_1 \rightarrow \mathbf{Lang}_2$  that commute with the source/target world linguistic indexings  $\mathbf{struc} \circ \mathbf{lang}_2 = \mathbf{lang}_1 \circ \mathbf{lang}$ . World morphisms are composable componentwise. Let **World** denote the category consisting of worlds and world morphisms.

**Polarity Principle:** *The linguistic indexing of the world is polar, having equivalent homogeneous and heterogeneous forms.*

We assume the linguistic indexing functor  $\mathbf{lang} : \mathbf{Struc} \rightarrow \mathbf{Lang}$  is a fibration (Cartesian passage). This gives a homogeneous representation of a world (all structures are in one category). By the equivalence between fibrations and index categories, we can alternately represent the world as the structure indexed category  $\mathbf{struc} : \mathbf{Lang}^{\text{op}} \rightarrow \mathbf{Cat}$  with the category of languages as its indexing category. This gives a heterogeneous representation of a world (structures are in many categories, each category indexed by the language of structures in that category). For any language  $\Sigma$ ,  $\mathbf{struc}(\Sigma)$  is the category whose objects are structures with underlying language  $\Sigma$ , and whose morphisms are structure morphisms whose underlying language morphism is the identity  $1_\Sigma$ . Here the indexed category  $\mathbf{struc}$  is the heterogenization of the fibration  $\mathbf{lang}$ , and  $\mathbf{lang}$  is (up to equivalence) the homogenization (fusion) of  $\mathbf{struc}$ . The indexed category of structures models structural heterogeneity, whereas the fibration of structures models structural homogeneity. In summary, a world can be represented as an indexed category  $\mathbf{Wrld} = \langle \mathbf{Lang}, \mathbf{struc} : \mathbf{Lang}^{\text{op}} \rightarrow \mathbf{Cat} \rangle$ .

By the equivalence between fibration morphisms and index morphisms, we can alternately represent a world morphism as an indexed morphism  $\langle \mathbf{lang}, \mathbf{struc} \rangle : \langle \mathbf{Lang}_1, \mathbf{struc}_1 \rangle \rightarrow \langle \mathbf{Lang}_2, \mathbf{struc}_2 \rangle$ , where  $\mathbf{lang} : \mathbf{Lang}_1 \rightarrow \mathbf{Lang}_2$  is the language functor and  $\mathbf{struc} : \mathbf{struc}_1 \Rightarrow \mathbf{lang}^{\text{op}} \circ \mathbf{struc}_2 : \mathbf{Lang}_1^{\text{op}} \rightarrow \mathbf{Cat}$  is a

structure natural transformation (bridge between passages) with  $\Sigma^{\text{th}}$  component functor  $\mathbf{struc}_\Sigma : \mathbf{struc}_1(\Sigma) \rightarrow \mathbf{struc}_2(\mathbf{lang}(\Sigma))$  for each language  $\Sigma$  in  $\mathbf{Lang}_1$ , where the naturality diagram  $\mathbf{struc}_1(\sigma) \circ \mathbf{struc}_\Sigma \cong \mathbf{struc}_{\Sigma'} \circ \mathbf{struc}_2(\mathbf{lang}(\sigma))$  connecting the source/target structure indexes, holds up to isomorphism for each language morphism  $\sigma : \Sigma \rightarrow \Sigma'$  in  $\mathbf{Lang}_1$ .

The structure indexed category has an underlying indexed set  $|\mathbf{struc}| = \mathbf{struc} \circ |-| : \mathbf{Lang}^{\text{op}} \rightarrow \mathbf{Set}$  obtained by forgetting morphism information; any language  $\Sigma$  is mapped to the set  $|\mathbf{struc}|(\Sigma)$  of all  $\Sigma$ -structures and any language morphism  $\sigma : \Sigma_1 \rightarrow \Sigma_2$  is mapped to the structure reduct  $|\mathbf{struc}|(\sigma) : |\mathbf{struc}|(\Sigma_2) \rightarrow |\mathbf{struc}|(\Sigma_1)$ .

## 2.2 Logical Expression

**Logic Principle:** *The description of the world involves semantics. It is logically meaningful, being based upon satisfaction.*

We assume (syntax) that any structure can be described using logical expressions (sentences) built from its language. In particular, we assume that there is a sentence functor (dual indexed set)  $\mathbf{sen} : \mathbf{Lang} \rightarrow \mathbf{Set}$  with the category of languages as its source (indexing category); any language  $\Sigma$  is mapped to the set  $\mathbf{sen}(\Sigma)$  of all sentences built upon it and any language morphism  $\sigma : \Sigma_1 \rightarrow \Sigma_2$  is mapped to the sentence translation map  $\mathbf{sen}(\sigma) : \mathbf{sen}(\Sigma_1) \rightarrow \mathbf{sen}(\Sigma_2)$  built upon it. Furthermore, we assume (semantics) that the indexed sets of structures and sentences are linked by satisfaction; any language  $\Sigma$  is mapped to a satisfaction relation (truth classification)  $\models_\Sigma \subseteq |\mathbf{struc}|(\Sigma) \times \mathbf{sen}(\Sigma)$ , where the symbolism  $M \models_\Sigma s$  expressing the assertion that ‘ $M$  satisfies  $s$ ’ means that  $\Sigma$ -sentence  $s$  is true when interpreted in  $\Sigma$ -structure  $M$ . A  $\Sigma$ -sentence  $s \in \mathbf{sen}(\Sigma)$  is a constraint (or theorem) of the  $\Sigma$ -structure  $M$ , and is said to be valid in  $M$ , when  $M$  satisfies  $s$ .

We also assume that satisfaction is preserved under sentence translation (infomorphism condition):  $|\mathbf{struc}|(\Sigma)(M_2) \models_{\Sigma_1} s_1$  *iff*  $M_2 \models_{\Sigma_2} \mathbf{sen}(\Sigma)(s_1)$  for target structure  $M_2 \in |\mathbf{struc}|(\Sigma_2)$  and source sentence  $s_1 \in \mathbf{sen}(\Sigma_1)$ . This expresses the invariance of truth under change of notation. Finally, the structure and sentence indexed sets can be combined with the satisfaction relations into a functor  $\mathbf{cls} : \mathbf{Lang} \rightarrow \mathbf{Cls}$  from the category of languages to the category of classifications and infomorphisms [7], whose composition with instance/type component functors gives the structure/sentence indexed sets  $\mathbf{cls}^{\text{op}} \circ \mathbf{inst} = |\mathbf{struc}|$  and  $\mathbf{cls} \circ \mathbf{typ} = \mathbf{sen}$ . We can think of the classification functor  $\mathbf{cls} : \mathbf{Lang} \rightarrow \mathbf{Cls}$  as a diagram within the ambient category of classifications and infomorphisms, indexed by languages and language morphisms. When composed with the lift functor (see below) logical expression and semantics extends to theories  $\mathbf{cls} \circ \mathbf{lift} : \mathbf{Lang} \rightarrow \mathbf{Cls}$ .

**Satisfaction Principle:** *The meaning of the world crucially involves both types and their particulars in classifications and infomorphisms. (This is the transfer to satisfaction of the second principle of Information Flow [1].)*

This principle motivates the use of structures (via satisfaction) as the interpretative objects (for the local logics that incorporate the regularities of a distributed system) and the use of structure morphisms (via truth invariance) as the interpretative morphisms (for the morphisms of local logics that incorporate the information flow of regularities of a distributed system).

### 2.3 Core Heterogeneity

**Core Principle:** *The architecture of the world description is concentrated in a 2-dimensional diagram of core indexed categories.*

Being a classification relation, satisfaction induces order. The intent of a structure  $M \in |\mathbf{struc}|(\Sigma)$  is the theory  $M^\Sigma = \{s \in \mathbf{sen}(\Sigma) \mid M \models_\Sigma s\}$ , the set of all sentences that are satisfied by  $M$ ; that is, the set of all theorems of  $M$ . This is called the (maximal) theory of the structure  $M$  and denoted be  $\mathbf{max}(M) = M^\Sigma$ . Structures are ordered by intent: two structures  $M_1, M_2 \in |\mathbf{struc}|(\Sigma)$  are ordered  $M_1 \leq_\Sigma M_2$  when  $M_1^\Sigma \supseteq M_2^\Sigma$ ; that is, when  $M_2 \models_\Sigma s$  implies  $M_1 \models_\Sigma s$  for all sentences  $s \in \mathbf{sen}(\Sigma)$ . This implies the condition, if  $M_1 \leq_\Sigma M_2$  and  $M_2 \models_\Sigma s$  then  $M_1 \models_\Sigma s$ , a bimodular condition stating that satisfaction respects the order on structures. Since preservation of satisfaction forms an infomorphism out of the structure and sentence maps, the structure map lifts to a monotonic function:  $\mathbf{struc}(\sigma) : \langle \mathbf{struc}(\Sigma_2), \leq_{\Sigma_2} \rangle \rightarrow \langle \mathbf{struc}(\Sigma_1), \leq_{\Sigma_1} \rangle$  for any language morphism  $\sigma : \Sigma_1 \rightarrow \Sigma_2$ . This means that the structure functor lifts to an indexed preorder (hence, category)  $\mathbf{struc}^\flat : \mathbf{Lang}^{\text{op}} \rightarrow \mathbf{Pre} \subseteq \mathbf{Cat}$  with  $\mathbf{struc}^\flat \circ |-| = |\mathbf{struc}|$ .

For any language  $\Sigma$ , a theory is a subset of sentences  $T \in \wp \mathbf{sen}(\Sigma)$ . Satisfaction lifts to theories: a structure satisfies a theory  $M \models_\Sigma T$  when it satisfies every sentence in it; then  $T$  is said to be valid in  $M$ . A theory  $T$  entails a sentence  $s$ ,  $T \vdash_\Sigma s$ , when any structure that satisfies  $T$  also satisfies  $s$ . The extent of a theory  $T^\Sigma$  is the set of all structures that satisfy it. The closure of a theory is the set of all entailed sentences — the intent of the extent  $T^\bullet = T^{\Sigma\Sigma}$ . Theories are ordered (entailment) by extent: two theories  $T_1, T_2 \in \wp \mathbf{sen}(\Sigma)$  are ordered  $T_1 \leq_\Sigma T_2$  when  $T_1^\Sigma \subseteq T_2^\Sigma$ . This implies the bimodular condition, if  $M \models_\Sigma T_1$  and  $T_1 \leq_\Sigma T_2$  then  $M \models_\Sigma T_2$ , which states that satisfaction respects entailment order. The sentence functor extends to theories in two adjoint ways. There is a direct image functor (dual indexed preorder)  $\mathbf{dir} : \mathbf{Lang} \rightarrow \mathbf{Pre}$ ; any language  $\Sigma$  is mapped to the entailment preorder  $\mathbf{th}(\Sigma) = \langle \wp \mathbf{sen}(\Sigma), \leq_\Sigma \rangle$  of all theories built upon it, and any language morphism  $\sigma : \Sigma_1 \rightarrow \Sigma_2$  is mapped to the theory direct image monotonic function  $\mathbf{dir}(\sigma) : \mathbf{th}(\Sigma_1) \rightarrow \mathbf{th}(\Sigma_2)$ ; there is a related functor  $\mathbf{lift} : \mathbf{Cls} \rightarrow \mathbf{Cls}$  that lifts the direct image functor on  $\mathbf{Set}$ . There is an inverse image functor (indexed preorder)  $\mathbf{inv} : \mathbf{Lang}^{\text{op}} \rightarrow \mathbf{Pre}$ ; any language  $\Sigma$  is mapped to the entailment preorder  $\mathbf{th}(\Sigma)$ , and any language morphism  $\sigma : \Sigma_1 \rightarrow \Sigma_2$  is mapped to the theory inverse image monotonic function  $\mathbf{inv}(\sigma) : \mathbf{th}(\Sigma_2) \rightarrow \mathbf{th}(\Sigma_1)$ .

Structures and theories can be embedded monotonically as concepts in the concept lattice of satisfaction [2]. Due to these embeddings, the intent order on

structures and the entailment order on theories are special cases of the lattice order of the satisfaction concept lattice. The concept lattice order is a generalization-specialization order with the more general concepts above and the more special concepts below. The fiber order of structures (theories) is the opposite of the intent (entailment) order on structures (theories) induced by the satisfaction concept lattice. Composition with the opposite preorder involution  $\alpha : \mathbf{Pre} \rightarrow \mathbf{Pre}$  gives the structure involuted indexed preorder  $\mathbf{struc}^b = \mathbf{struc}^b \circ \alpha : \mathbf{Lang}^{\text{op}} \rightarrow \mathbf{Pre}$  and the theory inverse image involuted indexed preorder  $\mathbf{inv}^\alpha = \mathbf{inv} \circ \alpha : \mathbf{Lang}^{\text{op}} \rightarrow \mathbf{Pre}$ . The structure involuted indexed preorder gives a heterogeneous representation of a flat world.

## 2.4 Flat Structures

Associated (homogenization) with the structure involuted indexed preorder  $\mathbf{struc}^b$  is the (flattened) structure fibration  $\mathbf{lang}^b : \mathbf{Struc}^b \rightarrow \mathbf{Lang}$  define as follows.  $\mathbf{Struc}^b$ , the category of (flat) structures and (flat) structure morphisms, is the Grothendieck construction of  $\mathbf{struc}^b$ . A (flat) structure  $\langle \Sigma, M \rangle$  consists of a language  $\Sigma$  and a  $\Sigma$ -structure  $M \in \mathbf{struc}(\Sigma)$  (so that  $\mathbf{lang}(M) = \Sigma$ ). A (flat) structure morphism  $\sigma : \langle \Sigma_1, M_1 \rangle \rightarrow \langle \Sigma_2, M_2 \rangle$  is a language morphism  $\sigma : \Sigma_1 \rightarrow \Sigma_2$  that preserves constraints (theorems):  $M_1 \models_{\Sigma_1} s_1$  implies  $M_2 \models_{\Sigma_2} \mathbf{sen}(\sigma)(s_1)$ . Equivalently, a (flat) structure morphism is an language morphism whose structure component maps the target structure to a specialization of the source structure  $\mathbf{struc}(\sigma)(M_2) \leq_{\Sigma_1} M_1$  or (intentwise, using theories)  $\mathbf{struc}(\sigma)(M_2)^{\Sigma_1} \leq_{\Sigma_1} M_1^{\Sigma_1}$ . Equivalently, a (flat) structure morphism is an language morphism whose inverse image sentence component maps the target structure intent to a specialization of the source structure intent  $\mathbf{sen}(\sigma)^{-1}(M_2^{\Sigma_2}) \leq_{\Sigma_1} M_1^{\Sigma_1}$ . This gives a homogeneous representation of a flat world, the fiber flattening of the category  $\mathbf{Struc}$ . The (flattened) language functor  $\mathbf{lang}^b$  is the projection, which maps an object  $\langle \Sigma, M \rangle$  to its indexing language  $\Sigma$  and maps a morphism  $\sigma : \langle \Sigma_1, M_1 \rangle \rightarrow \langle \Sigma_2, M_2 \rangle$  to its indexing language morphism  $\sigma : \Sigma_1 \rightarrow \Sigma_2$ . This is a fibration. For any language  $\Sigma$ , the identity language morphism  $1_\Sigma : \Sigma \rightarrow \Sigma$  is a (flat) structure morphism  $1_\Sigma : \langle \Sigma, M_1 \rangle \rightarrow \langle \Sigma, M_2 \rangle$  iff  $M_1 \geq_\Sigma M_2$ .

**Bimodular Principle:** *The description of the world factors through flat structures. Satisfaction is bimodular; that is, satisfaction respects structure morphisms.*

We assume that the language fibration factors  $\mathbf{lang} = \mathbf{flat} \circ \mathbf{lang}^b$  through the flattened language fibration  $\mathbf{lang}^b : \mathbf{Struc}^b \rightarrow \mathbf{Lang}$  by way of a structure flattening functor  $\mathbf{flat} : \mathbf{Struc} \rightarrow \mathbf{Struc}^b$ . This means the following: For any structure  $M$ , if  $\mathbf{lang}(M) = \Sigma$  then  $\mathbf{flat}(M) = \langle \Sigma, M \rangle$ . For any structure morphism  $f : M_1 \rightarrow M_2$ , if  $\mathbf{lang}(f : M_1 \rightarrow M_2) = \sigma : \Sigma_1 \rightarrow \Sigma_2$ , then  $\mathbf{flat}(f : M_1 \rightarrow M_2) = \sigma : \langle \Sigma_1, M_1 \rangle \rightarrow \langle \Sigma_2, M_2 \rangle$ . Hence,  $M_1 \geq_{\Sigma_1} \mathbf{struc}(\sigma)(M_2)$ ; equivalently,  $M_1 \models_{\Sigma_1} s_1$  implies  $M_2 \models_{\Sigma_2} \mathbf{sen}(\sigma)(s_1)$  (or  $\mathbf{struc}(\sigma)(M_2) \models_{\Sigma_1} s_1$ ) for every sentence  $s_1 \in \mathbf{sen}(\Sigma_1)$ . This implies the condition, if  $f : M_1 \rightarrow M_2$  is

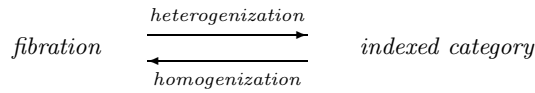
a structure morphism and  $M_1 \models_{\Sigma_1} s_1$  then  $M_2 \models_{\Sigma_2} \mathbf{sen}(\sigma)(s_1)$ , a bimodular condition stating that satisfaction respects structure morphisms. In particular, if  $f : M_1 \rightarrow M_2$  is a vertical structure morphism over language  $\Sigma$ , then  $M_1 \models_{\Sigma} s$  implies  $M_2 \models_{\Sigma} s$  for any sentence  $s \in \mathbf{sen}(\Sigma)$ ; that is,  $M_1 \geq_{\Sigma} M_2$ . Hence, the (flat) structure fiber over  $\Sigma$  is the underlying preorder of the structure fiber.

## 2.5 Theories

Associated (homogenization) with the theory inverse image involuted indexed preorder  $\mathbf{inv}^{\times}$  is the theory fibration  $\mathbf{lang} : \mathbf{Th} \rightarrow \mathbf{Lang}$  define as follows.  $\mathbf{Th}$ , the category of theories and theory morphisms, is the Grothendieck construction of  $\mathbf{inv}^{\times}$ . A theory  $\langle \Sigma, T \rangle$  consists of a language  $\Sigma$  and a  $\Sigma$ -theory  $T \in \wp \mathbf{sen}(\Sigma)$ . A theory morphism  $\sigma : \langle \Sigma_1, T_1 \rangle \rightarrow \langle \Sigma_2, T_2 \rangle$  is a language morphism  $\sigma : \Sigma_1 \rightarrow \Sigma_2$  that maps the target theory to a specialization of the source theory  $\mathbf{inv}(\sigma)(T_2) = \mathbf{sen}(\sigma)^{-1}(T_2^{\bullet}) \leq_{\Sigma_1} T_1$  iff  $\mathbf{sen}(\sigma)^{-1}(T_2^{\bullet}) \supseteq T_1$  iff  $\mathbf{sen}(\sigma)^{-1}(T_2^{\bullet}) \supseteq T_1^{\bullet}$ ; or that preserves entailment,  $T_1 \vdash_{\Sigma_1} t_1$  implies  $T_2 \vdash_{\Sigma_2} \mathbf{sen}(\sigma)(t_1)$  for any  $t_1 \in \mathbf{sen}(\Sigma_1)$ . Equivalently, a theory morphism is an language morphism that maps the source theory to a generalization of the target theory  $T_2 \leq_{\Sigma_2} \mathbf{dir}(\sigma)(T_1) = \wp \mathbf{sen}(\sigma)(T_1)$  iff  $T_2^{\bullet} \supseteq \wp \mathbf{sen}(\sigma)(T_1)$ . The projection fibration  $\mathbf{lang} : \mathbf{Th} \rightarrow \mathbf{Cls}$  maps a theory  $\langle \Sigma, T \rangle$  to the language  $\Sigma$  and maps a theory morphism  $\sigma : \langle \Sigma_1, T_1 \rangle \rightarrow \langle \Sigma_2, T_2 \rangle$  to the language morphism  $\sigma : \Sigma_1 \rightarrow \Sigma_2$ . For any language  $\Sigma$ , the identity language morphism  $1_{\Sigma} : \Sigma \rightarrow \Sigma$  is a (vertical) theory morphism  $1_{\Sigma} : \langle \Sigma, T_1 \rangle \rightarrow \langle \Sigma, T_2 \rangle$  iff  $T_1 \geq_A T_2$ . Hence, the theory fiber at language  $\Sigma$  is the opposite of the entailment theory preorder.

## 2.6 Logical Environments

In summary, a *logical environment* (Figure 1) is a more semantic version of an institution. It has both semantical and logical aspects. The semantical aspect is represented by a category of structures  $\mathbf{Struc}$  (the world) with a fibration  $\mathbf{lang} : \mathbf{Struc} \rightarrow \mathbf{Lang}$  (the world description) from structures into a category of logical languages  $\mathbf{Lang}$ ; or equivalently, an indexed category of structures  $\mathbf{struc} : \mathbf{Lang}^{\text{op}} \rightarrow \mathbf{Cat}$  with underlying indexed set  $|\mathbf{struc}| = \mathbf{struc} \circ |-| : \mathbf{Lang}^{\text{op}} \rightarrow \mathbf{Set}$ . The logical aspect is represented by a dual indexed set of sentences  $\mathbf{sen} : \mathbf{Lang} \rightarrow \mathbf{Set}$ . The semantical aspect is connected to the logical aspect via a functor  $\mathbf{cls} : \mathbf{Lang} \rightarrow \mathbf{Cls}$ , with instance projection  $\mathbf{cls}^{\text{op}} \circ \mathbf{inst} = |\mathbf{struc}|$  and type projection  $\mathbf{cls} \circ \mathbf{typ} = \mathbf{sen}$ . Structures and sentences are linked by satisfaction  $M \models_{\Sigma} s$ , with truth preserved under change of notation  $\mathbf{struc}(M_2) \models_{\Sigma_1} s_1$  iff  $M_2 \models_{\Sigma_1} \mathbf{sen}(s_1)$ . The concept order of satisfaction extentionally lifts to the (flat) structure indexed preorder  $\mathbf{struc}^b : \mathbf{Lang}^{\text{op}} \rightarrow \mathbf{Pre}$  and intentionally lifts to the theory direct image dual indexed preorder  $\mathbf{dir} : \mathbf{Lang} \rightarrow \mathbf{Pre}$  and (adjointly) to the theory inverse image indexed preorder  $\mathbf{inv} : \mathbf{Lang}^{\text{op}} \rightarrow \mathbf{Pre}$ . The (flat) structure indexed preorder is equivalent (homogenization  $\leftrightarrow$  heterogenization) to the (flattened) structure fibration (the flattened world)  $\mathbf{lang}^b : \mathbf{Struc}^b \rightarrow \mathbf{Lang}$ , such that description



**Fig. 1.** Logical Environment

factors through flat structures  $\mathbf{lang} = \mathbf{flat} \circ \mathbf{lang}^b$ . Either the theory direct image dual indexed preorder or the theory inverse image indexed preorder induce the theory fibration (the logical aspect)  $\mathbf{lang} : \mathbf{Th} \rightarrow \mathbf{Lang}$ . The crucial facts linking the semantical aspect of structures to the logical aspect of theories via satisfaction are (1) that the indexed set of structures is the instance projection of the classification functor and (2) that description factors through flat structures.

## 2.7 Examples

We use the three logical environments of equational logic **EQ**, first order logic **FOL** and information flow **IFC** as running examples. The world (structure category) of **EQ** consists of universal algebraic structures, that of **FOL** consists of first order logical structures, and that of **IFC** consists of classifications and infomorphisms.

*Example.* The logical environment IFC has  $\mathbf{Struc}_{\text{IFC}} = \mathbf{Cls}$  as its category of structures,  $\mathbf{Lang}_{\text{IFC}} = \mathbf{Set}$  as its category of languages, and the type functor  $\mathbf{lang}_{\text{IFC}} = \mathbf{typ} : \mathbf{Cls} \rightarrow \mathbf{Set}$  as its projection. For any set  $Y$ , the set  $\mathbf{struc}_{\text{IFC}}(Y)$  of  $Y$ -structures is the set of classifications with type set  $Y$ , the set  $\mathbf{sen}_{\text{IFC}}(Y)$  of  $Y$ -sentences is the set of  $Y$ -sequents, pairs  $\Gamma \vdash \Delta$  of subsets of types  $\Gamma, \Delta \subseteq Y$ , and a  $Y$ -classification  $\langle X, Y, \models \rangle$  satisfies a  $Y$ -sequent  $\Gamma \vdash \Delta$ , denoted  $\langle X, Y, \models \rangle \models_Y (\Gamma \vdash \Delta)$ , when for all instances  $x \in X$ ,  $x \models y$  for all  $y \in \Gamma$  implies  $x \models y'$  for some  $y' \in \Delta$ . For each function (language translation)  $f : Y \rightarrow Z$ , sentence translation along  $f$  is direct image squared on types  $\mathbf{sen}_{\text{IFC}}(f) : \mathbf{sen}_{\text{IFC}}(Y) \rightarrow \mathbf{sen}_{\text{IFC}}(Z) : (\Gamma \vdash \Delta) \mapsto (\wp f(\Gamma) \vdash \wp f(\Delta))$ , and structure translation along  $f$ ,  $\mathbf{struc}_{\text{IFC}}(f) : \mathbf{struc}_{\text{IFC}}(Z) \rightarrow \mathbf{struc}_{\text{IFC}}(Y)$ , maps a  $Z$ -classification  $C = \langle X, Z, \models \rangle$  to the  $Y$ -classification  $\mathbf{struc}_{\text{IFC}}(f)(C) = \langle X, Y, \models_f \rangle$ , where  $x \models_f y$  when  $x \models f(y)$ . The classification functor  $\mathbf{cls}_{\text{IFC}} : \mathbf{Set} \rightarrow \mathbf{Cls}$  maps a set  $Y$  to the classification

$\mathbf{cls}_{\text{IFC}}(Y) = \langle \mathbf{struc}_{\text{IFC}}(Y), \mathbf{sen}_{\text{IFC}}(Y), \models_Y \rangle$ , and maps a function  $f : Y \rightarrow Z$  to the infomorphism  $\mathbf{cls}_{\text{IFC}}(f) = \langle \mathbf{struc}_{\text{IFC}}(f), \mathbf{sen}_{\text{IFC}}(f) \rangle : \mathbf{cls}_{\text{IFC}}(Y) \rightrightarrows \mathbf{cls}_{\text{IFC}}(Z)$ . The logical environment IFC is a subenvironment of FOL when types are regarded as unary relation symbols.

*Example.* The logical system of equational logic (universal algebra) is represented by the logical environment EQ. The language category is  $\mathbf{Lang}_{\text{eq}} = \mathbf{Set}^{\mathbb{N}}$ , the  $\mathbb{N}^{\text{th}}$  power of  $\mathbf{Set}$ . A language  $\Phi$  is a family  $\Phi = \{\Phi_n \mid n \in \mathbb{N}\}$  of sets of function symbols, and a language morphism  $\phi : \Phi \rightarrow \Phi'$  is a family  $\{\phi_n : \Phi_n \rightarrow \Phi'_n \mid n \in \mathbb{N}\}$  of arity-preserving maps of function symbols. For any language  $\Phi$ , the set  $\mathbf{sen}_{\text{eq}}(\Phi)$  is the set of equations between  $\Phi$ -terms of function symbols. For any language morphism  $\phi : \Phi \rightarrow \Phi'$ , the sentence translation function  $\mathbf{sen}_{\text{eq}}(\phi) : \mathbf{sen}_{\text{eq}}(\Phi) \rightarrow \mathbf{sen}_{\text{eq}}(\Phi')$  is defined by function symbol substitution. A  $\Phi$ -structure  $A \in \mathbf{struc}_{\text{eq}}(\Phi)$  is a  $\Phi$ -algebra, consisting of a set (universe)  $A$  and a function (operation)  $f_A : A^n \rightarrow A$  for each function symbol  $f \in \Phi_n$ . A  $\Phi$ -structure morphism in  $\mathbf{struc}_{\text{eq}}(\Phi)$  is a  $\Phi$ -algebra morphism, consisting of a function (between universes)  $a : A \rightarrow A'$  that preserves operations  $f_A \cdot a = a^n \cdot f_{A'}$  for each function symbol  $f \in \Phi_n$ . Structure translation is reduct with symbol translation. Satisfaction is as usual.

*Example.* The logical system of unsorted first-order logic with equality is represented by the logical environment FOL. This extends the logical environment of equational logic by adding relation symbols. The language category is  $\mathbf{Lang}_{\text{fol}} = \mathbf{Set}^{\mathbb{N}} \times \mathbf{Set}^{\mathbb{N}} \cong \mathbf{Set}^{\mathbb{N} + \mathbb{N}} \cong \mathbf{Set}^{\mathbb{N}}$ , the square of the  $\mathbb{N}^{\text{th}}$  power of  $\mathbf{Set}$ . A language  $\langle \Phi, \Psi \rangle$  is a family  $\Phi$  as above, plus a family  $\Psi = \{\Psi_n \mid n \in \mathbb{N}\}$  of sets of relation symbols of arity  $n$ . A language morphism  $\langle \phi, \psi \rangle : \langle \Phi, \Psi \rangle \rightarrow \langle \Phi', \Psi' \rangle$  is a family  $\phi$  as above, plus a family  $\psi : \Psi \rightarrow \Psi'$  of arity-preserving maps of relation symbols. Sentences are the usual first order sentences. For any language  $\langle \Phi, \Psi \rangle$ , the set  $\mathbf{sen}_{\text{fol}}(\Phi, \Psi)$  of  $\langle \Phi, \Psi \rangle$ -sentences consists of closed first-order formulae using function symbols from  $\Phi$  and relation symbols from  $\Psi$ . For any language morphism  $\langle \phi, \psi \rangle : \langle \Phi, \Psi \rangle \rightarrow \langle \Phi', \Psi' \rangle$ , the sentence translation function  $\mathbf{sen}_{\text{fol}}(\phi, \psi) : \mathbf{sen}_{\text{fol}}(\Phi, \Psi) \rightarrow \mathbf{sen}_{\text{fol}}(\Phi', \Psi')$  is defined by symbol substitution. A  $\langle \Phi, \Psi \rangle$ -structure  $A \in \mathbf{struc}_{\text{fol}}(\Phi, \Psi)$  is a  $\Phi$ -algebra  $A$  (as above) and a subset  $R_A \subseteq A^n$  for each relation symbol  $R \in \Psi_n$ . A  $\langle \Phi, \Psi \rangle$ -structure morphism in  $\mathbf{struc}_{\text{fol}}(\Phi, \Psi)$  is a  $\Phi$ -algebra morphism (as above), which preserves relations  $\wp a^n(R_A) \subseteq R_{A'}$  for each relation symbol  $R \in \Psi_n$ . Structure translation is reduct with symbol translation. Satisfaction is as usual. The institution FOL can be extended to the institution  $\mathbf{FOL}^*$ , which replaces language maps with language interpretations  $\langle \phi, \psi \rangle : \langle \Phi, \Psi \rangle \rightarrow \langle \Phi', \Psi' \rangle$  mapping function symbols to terms of the same arity  $\{\phi_n : \Phi_n \rightarrow \mathbf{term}(\Phi')_n \mid n \in \mathbb{N}\}$  and mapping relation symbols to expressions of the same arity  $\{\psi_n : \Psi_n \rightarrow \mathbf{expr}(\Psi')_n \mid n \in \mathbb{N}\}$ .

*Example.* The category  $\mathbf{Set}_{\subseteq}$  has subsets  $Y \subseteq X$  as objects and restrictions ( $g \subseteq f$ ) :  $(Y_1 \subseteq X_1) \rightarrow (Y_2 \subseteq X_2)$  as morphisms, where  $f : X_1 \rightarrow X_2$  is a function and  $g : Y_1 \rightarrow Y_2$  is a restriction of  $f$ . The category  $\mathbf{Cls}_{\subseteq}$  has  $\mathbf{Set}_{\subseteq}$  as its component instance category and  $\mathbf{Set}$  as its component type category. An object  $\langle Y, A \rangle$



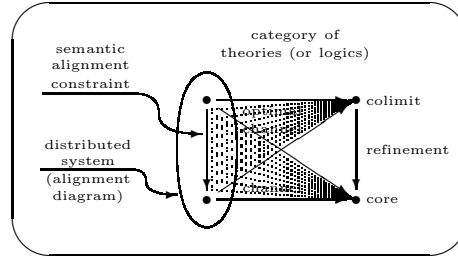
in  $\mathbf{Cls}_\subseteq$  consists of a classification  $A$  and a subset of instances  $Y \subseteq \mathbf{inst}(A)$ . A morphism  $\langle g, f \rangle : \langle Y_1, A_1 \rangle \Rightarrow \langle Y_2, A_2 \rangle$  in  $\mathbf{Cls}_\subseteq$  consists of an infomorphism  $f : A_1 \Rightarrow A_2$  and a restriction  $g : Y_2 \rightarrow Y_1$  of the instance function  $\mathbf{inst}(f)$ . The logical environment  $\mathbf{IFS}$  has  $\mathbf{Lang}_{\mathbf{IFS}} = \mathbf{Set}$  as its category of languages,  $\mathbf{Struc}_{\mathbf{IFS}} = \mathbf{Cls}_\subseteq$  as its category of structures, and  $\mathbf{lang}_{\mathbf{IFS}} = \mathbf{typ} : \mathbf{Cls}_\subseteq \rightarrow \mathbf{Set}$  as its language index functor. This logical environment allows the definition of a normal subset of instances.

### 3 Information Flow

#### 3.1 Distributed Systems

**System Principle:** *Information flow results from regularities in a distributed system.* (This is the first principle of Information Flow [1].)

This principle motivates the representation of distributed systems by diagrams of objects that can incorporate regularities. Eventually, we will argue that these objects should be local logics.



**Fig. 2.** Distributed System

The semantic integration of ontologies [7], [4] can be represented by alignment and unification (Figure 2): aligning a distributed system of ontologies by building a suitable diagram of logics and unifying the distributed system of ontologies along a channel covering the underlying diagram of structures. The logics in the alignment diagram represent the individual ontologies, and the morphisms between logics in the alignment diagram represent the semantic alignment constraints. An example of semantic alignment constraints is the representation of an equivalent pair of types in two ontologies being aligned by a single type in a mediating ontology, with two mappings from this mediating type back to the equivalent pair of types. The alignment diagram represents a semantically constrained distributed system of ontologies, with individual logics representing parts of the system. Any covering channel over the underlying diagram of structures has a core that represents the whole system in some respect. Unification

forms a covering channel of logics that connects the distributed system to the fusion logic — the meet, in the logic fiber over the underlying core, of the direct image of the diagram of logics along the underlying channel of structures.

**Structure Principle:** *Information flow crucially involves structures of the world.* (This is the second principle of Information Flow [1], abstracted from classifications to structures. A classification is just one example of a structure.)

This principle motivates the use of structures as the indexing objects for the (local) logics that incorporate the regularities of a distributed system and the use of structure morphisms as the indexing links for the morphisms of (local) logics that incorporate the information flow of regularities of a distributed system.

A distributed system  $\mathcal{A} : \mathbf{I} \rightarrow \mathbf{Struc}$  consists of an indexed family  $\{\mathcal{A}_i \mid i \in |\mathbf{I}|\}$  of structures together with an indexed family  $\{\mathcal{A}_e : A_i \rightarrow A_j \mid (e : i \rightarrow j) \in \mathbf{I}\}$  of structure morphisms; that is, a distributed system is a diagram in the structure category **Struc**. We think of the component structures  $\mathcal{A}_i$  as being parts of the system. We would also like to represent the whole system as a structure, where we might have different representative structures for different purposes. The theory of part-whole relations is called mereology. It studies how parts are related to wholes, and how parts are related to other parts within a whole. In a distributed system, the part to part relationships are modeled by the structure morphisms  $\mathcal{A}_e : A_i \rightarrow A_j$ . In Information Flow, we can model the whole as a structure  $C$  and model the part-whole relationship between some part  $A$  and the whole with a structure morphism  $g : A \rightarrow C$ .

### 3.2 Information Channels

An information channel  $\mathcal{C} : A \xrightarrow{g} \Delta(C)$  over a world (category of structures) **Struc** consists of an  $I$ -indexed family  $\mathcal{C} = \{g_i : A_i \rightarrow C\}_{i \in I}$  of structure morphisms with a common target structure  $C$ , called the core of the channel. A channel  $\mathcal{C} : A \xrightarrow{g} \Delta(C)$  covers a distributed system  $\mathcal{A} : \mathbf{I} \rightarrow \mathbf{Struc}$  when  $I = |\mathbf{I}|$  and the channel component morphisms commute with the distributed system morphisms  $A_i \xrightarrow{g_i} C = A_i \xrightarrow{\mathcal{A}_e} A_j \xrightarrow{g_j} C$  for  $e : i \rightarrow j$  in  $\mathbf{I}$ . A covering channel  $\mathcal{C} : \mathcal{A} \xrightarrow{g} \Delta(C)$  is essentially a cocone over diagram  $\mathcal{A}$ . For any two covering channels  $\mathcal{C} : \mathcal{A} \xrightarrow{g} \Delta(C)$  and  $\mathcal{D} : \mathcal{A} \xrightarrow{h} \Delta(D)$  over the same distributed system  $\mathcal{A}$ , a refinement (mediating morphism) is a structure morphism between cores  $r : C \rightarrow D$  that commutes with the channel component morphisms  $A_i \xrightarrow{g_i} C \xrightarrow{r} D = A_i \xrightarrow{h_j} C$  for  $i \in |\mathbf{I}|$ . A channel  $\mathcal{C}_{\text{opt}} : \mathcal{A} \xrightarrow{k} \Delta(\coprod \mathcal{A})$  is a minimal cover of a distributed system  $\mathcal{A}$  when it covers  $\mathcal{A}$  and for any other covering channel  $\mathcal{D}$  there is a unique refinement  $[\mathcal{D}] : \coprod \mathcal{A} \rightarrow D$  from  $\mathcal{C}_{\text{opt}}$  to  $\mathcal{D}$ . A minimal cover is essentially a colimiting cocone over diagram  $\mathcal{A}$ . Any two minimal covers are isomorphic.

Information flow has two concerns with respect to channels: (1) given a distributed system and some viewpoint (scientific, technological, social, etc.), how should the whole system be modeled; and (2) how does the natural logic of one

component part of a system affect another component part. The first concern, realizing a channel core, can have several solutions. An optimal solution, the colimit, is discussed in the section on cocompleteness and cocontinuity. The second concern, involving distributed inference rules, local logics and information flow, is discussed in other succeeding sections.

**Connection Principle:** *It is by virtue of regularities among connections that information about some components of a distributed system carries information about other components.* (This is the third principle of Information Flow [1].)

This principle motivates the use of logics over structures, which lift theories over languages, to represent information flow over covering channels of a distributed system. For a simple example of information flow, consider two component parts  $A_i \xrightarrow{g_i} C \xleftarrow{g_j} A_j$  with underlying language morphisms  $\gamma_i$  and  $\gamma_j$  that are connected to the core structure  $C$  by being essentially projections  $A_i \cong \mathbf{struc}(\gamma_i)(C)$  and  $\mathbf{struc}(\gamma_j)(C) \cong A_j$ . Then  $A_i$ 's satisfying sentence  $a_i$  carries the information that  $A_j$  satisfies sentence  $a_j$ , relative to the channel  $C$ , if the translation  $\mathbf{sen}(\gamma_i)(a_i)$  entails the translation  $\mathbf{sen}(\gamma_j)(a_j)$  in the theory  $\mathbf{max}(C)$ .

### 3.3 Inference Rules

In the section we paraphrase the discussion in the first part of [1]. To see how unsound and incomplete logics arise in reasoning about distributed systems, we consider the diagram

$$P \xrightarrow{p} C \xleftarrow{d} D$$

called a binary channel. This consists of a proximal structure  $P$  a distal structure  $D$  and a connecting structure  $C$ . We think of this binary channel as representing a distributed system having proximal part  $P$ , distal part  $D$  and whole (or core)  $C$ .

We are interested in discovering what kind of theory of the distal part is available to someone with complete knowledge of the proximal part. The diagram suggests breaking the problem up into two parts, the problem of going directly along  $p$  from proximal component  $P$  to core component  $C$ , and the problem of going inversely along  $d$  from core component  $C$  to distal component  $D$ . We can discuss both steps at once by considering a single structure morphism  $f : M_1 \rightarrow M_2$  (in the above,  $f$  can be either  $p : P \rightarrow C$  or  $d : C \rightarrow D$ ). Image someone who wants to reason about one side by using the induced theory of the other side.

Consider the following “rules of inference” along an structure morphism  $f : M_1 \rightarrow M_2$  with underlying language morphism  $\mathbf{lang}(f) = \sigma : \Sigma_1 \rightarrow \Sigma_2$ . The first says that from any source sentence  $s_1 \in \mathbf{sen}(\Sigma_1)$  we can infer the target sentence  $\mathbf{sen}(\sigma)(s_1) \in \mathbf{sen}(\Sigma_2)$ . The second is read similarly.

$$f\text{-Intro: } \frac{s_1}{\mathbf{sen}(\sigma)(s_1)} \qquad f\text{-Elim: } \frac{\mathbf{sen}(\sigma)(s_1)}{s_1}$$

The first rule allows us to move along the structure morphism  $f$  from a source sentence to a target sentence, whereas the second rule allows us to move in the opposite direction. These inference rules have very important properties. First consider the preservation of validity and nonvalidity.

A rule preserves validity when it leads from premise constraints to conclusion constraints. The  $f$ -**Intro** rule preserves validity: if the premise  $s_1$  is valid in  $M_1$ ,  $M_1 \models_{\Sigma_1} s_1$ , then the conclusion  $\mathbf{sen}(\sigma)(s_1)$  is valid in  $M_2$ ,  $M_2 \models_{\Sigma_2} \mathbf{sen}(\sigma)(s_1)$ . This follows immediately from the definition of structure morphism. The  $f$ -**Elim** rule does not preserve validity. It is possible to have a constraint  $\mathbf{sen}(\sigma)(s_1)$  of  $M_2$ ,  $M_2 \models_{\Sigma_2} \mathbf{sen}(\sigma)(s_1)$ , such that  $s_1$  has counterexample  $M_1$ ,  $M_1 \not\models_{\Sigma_1} s_1$ . However,  $M_1$  is not a counterexample of  $s_1$  if  $M_1$  is a specialization of  $\mathbf{struc}(\sigma)(M_2)$ ,  $M_1 \leq_{\Sigma_1} \mathbf{struc}(\sigma)(M_2)$ ; that is, the  $f$ -**Elim** rule is sound when  $M_1 \cong_{\Sigma_1} \mathbf{struc}(\sigma)(M_2)$ . In particular, the  $f$ -**Elim** rule preserves validity along structure-isomorphic structure morphisms.

A rule preserves nonvalidity when it leads from premise nonconstraints to conclusion nonconstraints. The  $f$ -**Elim** rule preserves nonvalidity: if the premise  $\mathbf{sen}(\sigma)(s_1)$  is not valid in  $M_2$ ,  $M_2 \not\models_{\Sigma_2} \mathbf{sen}(\sigma)(s_1)$ , then the conclusion  $s_1$  is not valid in  $M_1$ ,  $M_1 \not\models_{\Sigma_1} s_1$ . This also follows immediately from the definition of structure morphism. The  $f$ -**Intro** rule does not preserve nonvalidity. It is possible that  $s_1$  has counterexample  $M_1$ ,  $M_1 \not\models_{\Sigma_1} s_1$ , where  $\mathbf{sen}(\sigma)(s_1)$  is a constraint of  $M_2$ ,  $M_2 \models_{\Sigma_2} \mathbf{sen}(\sigma)(s_1)$ . However,  $\mathbf{sen}(\sigma)(s_1)$  is not a constraint of  $M_2$  if  $M_1$  is a specialization of  $\mathbf{struc}(\sigma)(M_2)$ ,  $M_1 \leq_{\Sigma_1} \mathbf{struc}(\sigma)(M_2)$ ; that is, the rule preserves nonvalidity when  $M_1 \cong_{\Sigma_1} \mathbf{struc}(\sigma)(M_2)$ . In particular, the  $f$ -**Intro** rule preserves nonvalidity along structure-isomorphic structure morphisms.

Summarizing the above, the rule of  $f$ -**Intro** preserves validity, but not nonvalidity; whereas, the rule of  $f$ -**Elim** preserves nonvalidity, but not validity. In terms of distributed systems, when using the  $f$ -**Intro** rule any constraint that holds for a component translates directly to a constraint about the whole system, and when using the  $f$ -**Elim** rule any constraint about the whole system translates inversely to a constraint of those parts that really are (up to isomorphism) a component of the system structure. Returning to the binary channel

$$P \xrightarrow{p} C \xleftarrow{d} D$$

depicted previously, we wanted to know what happens when we use the sound and complete theory  $\mathbf{max}(P) = \{s \in \mathbf{sen}(\Sigma_P) \mid P \models_{\Sigma_P} s\}$  of the proximal structure to reason about the distal structure. On the one hand, we have seen that  $p$ -**Intro** preserves validity, but not nonvalidity; so that the theory we obtain at  $C$  may be sound (any constraint of  $P$  maps to a constraint of  $C$ ), but not necessarily complete (we may have no sentence or only nonvalid sentences of  $P$  mapping to a particular constraint of  $C$  — there may be constraints of  $C$  that are missed). On the other hand, following  $p$ -**Intro** by  $d$ -**Elim** means that we lose our guarantee that the resulting distal theory is either sound or complete. A sentence about the distal structure obtained from a sentence about the proximal structure in this way is guaranteed to apply when the distal structure is connected (up to isomorphism) to the proximal structure in the channel.

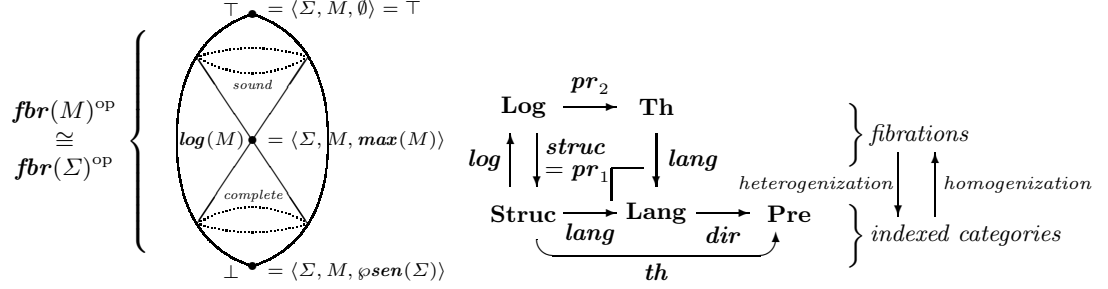


Fig. 3. The Logic Context

### 3.4 Local Logics

**Channel Principle:** *The regularities of a given distributed system are relative to its analysis in terms of information channels.* (This is the fourth principle of Information Flow [1].)

Paraphrasing and quoting [1] (Lecture 12), when “reasoning about a distributed system with component” parts of various kinds, the component parts will typically be describe in quite different ways with different languages. Along with these different languages “it is natural to think of each of the components as having its own logic”, expressed in its own language. “In this way, the distributed system” of structures “gives rise to a distributed system of local logics. The interactions of the local logics reflect the behavior of the system as a whole.” The concept of a (local) logic represents the regularities of a distributed system and tracks what happens when we reason at a distance.

A (local) logic  $L = \langle \Sigma, M, T \rangle$  consists of an indexing language  $\Sigma$ , a  $\Sigma$ -structure  $M \in \mathbf{struc}(\Sigma)$  and a  $\Sigma$ -theory  $T \in \mathbf{th}(\Sigma)$ . For any fixed structure  $M$  with underlying language  $\Sigma$ , the set of all logics with that structure is a preordered set under the theory order:  $\langle \Sigma, M, T_1 \rangle \leq \langle \Sigma, M, T_2 \rangle$  when  $T_1 \leq T_2$ . This is the (opposite of the) fiber over  $M$  with respect to the logic-to-structure projection functor  $\mathbf{struc}$ . There are larger fibers. For any fixed language  $\Sigma$ , the set of all logics with that language is a preordered set under the structure and theory orders:  $\langle \Sigma, M_1, T_1 \rangle \leq \langle \Sigma, M_2, T_2 \rangle$  when  $M_1 \leq_\Sigma M_2$  and  $T_1 \leq_\Sigma T_2$ . This is the (opposite of the) fiber over  $\Sigma$  with respect to the composite functor  $\mathbf{struc} \circ \mathbf{lang}$ . Any structure  $M$  with underlying language  $\Sigma$  induces the natural logic  $\mathbf{log}(M) = \langle \Sigma, M, \mathbf{max}(M) \rangle = \langle \Sigma, M, M^\Sigma \rangle$ . If two structures are ordered  $M_1 \leq_\Sigma M_2$ , then their logics are ordered  $\mathbf{log}(M_1) \leq_\Sigma \mathbf{log}(M_2)$ , since  $M_1 \leq_\Sigma M_2$  iff  $M_1^\Sigma \supseteq M_2^\Sigma$  iff  $M_1^\Sigma \leq_\Sigma M_2^\Sigma$ .

A logic morphism  $f : \langle \Sigma_1, M_1, T_1 \rangle \rightarrow \langle \Sigma_2, M_2, T_2 \rangle$  is a structure morphism  $f : M_1 \rightarrow M_2$ , whose underlying language morphism  $\sigma : \Sigma_1 \rightarrow \Sigma_2$  is also a theory morphism  $\sigma : \langle \Sigma_1, T_1 \rangle \rightarrow \langle \Sigma_2, T_2 \rangle$ . The category of logics  $\mathbf{Log}$  has logics as objects and logic morphisms as morphisms. It is describable (Figure 3) as either the pullback of the indexing functors for structures and theories, or the Grothendieck

construction of the dual indexed preorder  $\mathbf{th} = \mathbf{lang} \circ \mathbf{dir} : \mathbf{Struc} \rightarrow \mathbf{Pre}$ . The projection functors  $\mathbf{struc} = \mathbf{pr}_1 : \mathbf{Log} \rightarrow \mathbf{Struc}$  and  $\mathbf{pr}_2 : \mathbf{Log} \rightarrow \mathbf{Th}$  satisfy the pullback condition  $\mathbf{pr}_1 \circ \mathbf{lang} = \mathbf{pr}_2 \circ \mathbf{lang}$ . The first pullback projection is the fibration associated (homogenization) with the dual indexed preorder  $\mathbf{th} : \mathbf{Struc} \rightarrow \mathbf{Pre}$ . This pullback projection is the structure lift of the theory indexing functor  $\mathbf{lang} : \mathbf{Th} \rightarrow \mathbf{Lang}$ . The second pullback projection is the theory lift of the structure indexing functor  $\mathbf{lang} : \mathbf{Struc} \rightarrow \mathbf{Lang}$ . Any structure morphism  $f : M_1 \rightarrow M_2$ , with underlying language morphism  $\sigma : \Sigma_1 \rightarrow \Sigma_2$ , induces the logic morphism  $f : \mathbf{log}(M_1) \rightarrow \mathbf{log}(M_2)$  between natural logics. This is well-defined, since validity-preservation is equivalent to  $M_2^{\Sigma_2} \supseteq \wp \mathbf{sen}(\sigma)(M_1^{\Sigma_1})$ . Hence, there is a functor  $\mathbf{log} : \mathbf{Struc} \rightarrow \mathbf{Log}$  (Figure 3) satisfying  $\mathbf{log} \circ \mathbf{base} = 1_{\mathbf{Struc}}$ . A local logic  $\langle \Sigma, M, T \rangle$  in the fiber over  $M$  does not compare to the natural logic  $\mathbf{log}(M)$ , unless it is either sound or complete.

In general, logics may be neither sound nor (logically) complete. A logic  $\langle \Sigma, M, T \rangle$  is *sound* when the structure  $M$  satisfies the theory  $T$  ( $T$  is valid in  $M$ ),  $M \models_{\Sigma} T$ ; equivalently, when  $M^{\Sigma} \supseteq T$  or  $M^{\Sigma} \supseteq T^{\bullet}$  or  $M^{\Sigma} \leq_{\Sigma} T$ . A logic  $\langle \Sigma, M, T \rangle$  is *complete* when every sentence satisfied by  $M$  is entailed by  $T$ ; that is, when  $M \models_{\Sigma} t$  implies  $T \vdash_{\Sigma} t$  for all sentences  $t \in \mathbf{sen}(\Sigma)$ ; equivalently, when  $T^{\bullet} \supseteq M^{\Sigma}$  or  $T \leq_{\Sigma} M^{\Sigma}$ ; equivalently, when  $T \not\models_{\Sigma} t$  implies  $M \not\models_{\Sigma} t$  for all sentences  $t \in \mathbf{sen}(\Sigma)$ . A logic  $\langle \Sigma, M, T \rangle$  is sound (complete) iff the identity (flat) structure morphism  $1_M : M \rightarrow M$  with underlying identity language morphism  $1_{\Sigma} : \Sigma \rightarrow \Sigma$  is a logic morphism  $\eta_{\langle \Sigma, M, T \rangle} = 1_M : \langle \Sigma, M, T \rangle \rightarrow \mathbf{log}(M)$  ( $\varepsilon_{\langle \Sigma, M, T \rangle} = 1_M : \mathbf{log}(M) \rightarrow \langle \Sigma, M, T \rangle$ ). The only sound and complete logics are those equivalent to the natural logic  $\mathbf{log}(M)$  for some structure  $M$ . Sound logics form a reflective subcategory of all logics  $\mathbf{Snd} \subseteq \mathbf{Log}$  with unit natural transformation  $\eta : 1_{\mathbf{Snd}} \Rightarrow \mathbf{struc} \circ \mathbf{log}$ . Complete logics form a coreflective subcategory of all logics  $\mathbf{Cmp} \subseteq \mathbf{Log}$  with counit natural transformation  $\varepsilon : \mathbf{log} \circ \mathbf{struc} \Rightarrow 1_{\mathbf{Cmp}}$ .

### 3.5 Colimits

**Completeness/continuity Principle:** *The world is cocomplete and the description of the world is cocontinuous.*

We assume the logical environment (Figure 1) is cocomplete. This means that it has a cocomplete category of structures  $\mathbf{Struc}$ , a cocomplete category of logical languages  $\mathbf{Lang}$ , and a cocontinuous fibration  $\mathbf{lang} : \mathbf{Struc} \rightarrow \mathbf{Lang}$ . Hence, all colimits (universal constructions) of structures and languages are possible, and the underlying language of the colimit of a diagram of structures is the colimit of the underlying diagram of languages. All the examples of logical environments (IFC, EQ, FOL) are cocomplete.

In approach advocated here, unpopulated ontologies (no world information and no semantics) are represented by theories and the optimal semantic integration of unpopulated ontologies is represented by the colimit construction of theories, whereas populated ontologies (both world information and semantics) are represented by logics, and the optimal semantic integration of populated

ontologies is represented by the colimit construction of logics. Colimits in the category of theories (logics) can be used to fuse together smaller theories (logics) to form larger ones. The colimit construction in the category of theories (logics) forms an optimal channel  $\mathcal{C}_{\text{opt}} : \mathcal{A} \xrightarrow{\ell} \Delta(\coprod \mathcal{A})$ . The fusion theory (logic)  $\coprod \mathcal{A}$  is formed by information flow over the optimal channel: direct image flow followed by meet in the lattices of theories (logics). The colimit construction is based upon the colimit theorem, a powerful, general criterion for when such colimits of theories actually exist. It allows us to use for semantic integration, the same flow and lattice operators on logics-over-models as we do for theories-over-languages.

**Theorem 1.** (*Cocompleteness/Cocontinuity*) [5] *For any logical environment, the projection functors  $\mathbf{lang} : \mathbf{Struc}^b \rightarrow \mathbf{Lang}$  and  $\mathbf{lang} : \mathbf{Th} \rightarrow \mathbf{Lang}$  reflect colimits. Hence, if the logical environment is cocomplete, then its category  $\mathbf{Struc}^b$  of (flat) structures and structure morphisms and its category  $\mathbf{Th}$  of theories and theory morphisms are cocomplete and the projection functors are cocontinuous.*

When the structure category  $\mathbf{Struc}$  is cocomplete, information flow in a channel  $\mathcal{C} : \mathcal{A} \xrightarrow{g} \Delta(C)$  covering a distributed system  $\mathcal{A}$  can be factored through information flow in the optimal channel followed by direct image along the unique mediating structure morphism  $\coprod \mathcal{A} \xrightarrow{m} C$ .

### 3.6 Information Flow

Information flow over a channel is defined as a two-step process: direct image flow along the component structure morphisms of the channel, followed by the meet operation in the lattice of logics over the core.

Given a structure morphism  $f : M_1 \rightarrow M_2$  with underlying language morphism  $\sigma : \Sigma_1 \rightarrow \Sigma_2$  and a source logic  $L_1 = \langle \Sigma_1, M_1, T_1 \rangle$ , the *direct image* is the target logic  $\mathbf{dir}(f)(L_1) = \langle \Sigma_2, M_2, \wp \mathbf{sen}(\sigma)(T_1) \rangle$ . The direct image is the greatest logic  $L_2$  on the target structure  $M_2$  such that  $f$  is a logic morphism from  $L_1$  to  $L_2$ ; that is,  $f : L_1 \rightarrow \mathbf{dir}(f)(L_1)$  is a logic morphism; and if  $f : L_1 \rightarrow L_2$  is a logic morphism, then  $L_2 \leq \mathbf{dir}(f)(L_1)$ .

**Proposition 1.** *For any structure morphism  $f : M_1 \rightarrow M_2$ , direct image preserves soundness: if a source logic  $L_1$  is sound, then the direct image logic  $\mathbf{dir}(f)(L_1)$  is also sound.*

*Proof:* Assume the source logic  $\langle \Sigma_1, M_1, T_1 \rangle$  is sound. This means that  $M_1^{\Sigma_1} \leq_{\Sigma_1} T_1$  (sentences inside the closure  $T_1^\bullet$  are valid). Since  $f : M_1 \rightarrow M_2$  is an structure morphism, with the underlying language morphism  $\mathbf{struc}(f) = \sigma : \Sigma_1 \rightarrow \Sigma_2$ ,  $\mathbf{struc}(\sigma)(M_2)^{\Sigma_1} \leq_{\Sigma_1} M_1^{\Sigma_1}$ . Thus,  $\mathbf{struc}(\sigma)(M_2)^{\Sigma_1} \leq_{\Sigma_1} T_1$ . By satisfaction invariance,  $\mathbf{struc}(\sigma)(M_2)^{\Sigma_1} \leq_{\Sigma_1} T_1$  iff  $M_2^{\Sigma_2} \leq_{\Sigma_2} \wp \mathbf{sen}(\sigma)(T_1)$ , this means that the direct image  $\mathbf{dir}(f)(L_1) = \langle \Sigma_2, M_2, \wp \mathbf{sen}(\sigma)(T_1) \rangle$  is sound. ■

Given a structure morphism  $f : M_1 \rightarrow M_2$  with underlying language morphism  $\sigma : \Sigma_1 \rightarrow \Sigma_2$  and a target logic  $L_2 = \langle \Sigma_2, M_2, T_2 \rangle$ , the *inverse image* is

the source logic  $\mathbf{inv}(f)(L_2) = \langle \Sigma_1, M_1, \mathbf{sen}(\sigma)^{-1}(T_2^\bullet) \rangle$ , where  $\mathbf{sen}(\sigma)^{-1}(T_2^\bullet) = \{t_1 \in \mathbf{sen}(\Sigma_1) \mid T_2 \vdash_{\Sigma_2} \mathbf{sen}(\sigma)(t_1)\}$ . The inverse image is the least logic  $L_1$  on source structure  $M_1$  such that  $f$  is a logic morphism from  $L_1$  to  $L_2$ ; that is,  $f : \mathbf{inv}(f)(L_2) \rightarrow L_2$  is a logic morphism; and if  $f : L_1 \rightarrow L_2$  is a logic morphism, then  $\mathbf{inv}(f)(L_2) \leq L_1$ .

**Proposition 2.** *For any structure morphism  $f : M_1 \rightarrow M_2$ , inverse image preserves completeness: if a target logic  $L_2$  is complete, then the inverse image logic  $\mathbf{inv}(f)(L_2)$  is also complete.*

*Proof.* Assume target logic  $\langle \Sigma_2, M_2, T_2 \rangle$  is complete. This means that  $T_2^\bullet \leq_{\Sigma_2} M_2^{\Sigma_2}$ . Hence,  $\mathbf{sen}(\sigma)^{-1}(T_2^\bullet) \leq_{\Sigma_1} \mathbf{sen}(\sigma)^{-1}(M_2^{\Sigma_2})$ . Since  $f : M_1 \rightarrow M_2$  is an structure morphism with language morphism  $\mathbf{struc}(f) = \sigma : \Sigma_1 \rightarrow \Sigma_2$ , we have  $\mathbf{sen}(\sigma)^{-1}(M_2^{\Sigma_2}) \leq_{\Sigma_1} M_1^{\Sigma_1}$ . Hence,  $\mathbf{sen}(\sigma)^{-1}(T_2^\bullet) \leq_{\Sigma_1} M_1^{\Sigma_1}$ . This means that the inverse image  $\mathbf{inv}(f)(L_2) = \langle \Sigma_1, M_1, \mathbf{sen}(\sigma)^{-1}(T_2^\bullet) \rangle$  is complete. ■

## 4 Conclusion

We have defined logical environments, semantic versions of institutions, and have demonstrated how important concepts in IF theory, such as distributed systems, channels and information flow, can be defined within logical environments. Thus, IF theory abstracts and extends INS theory.

## References

1. Barwise, J., Seligman, J.: Information Flow: The Logic of Distributed Systems. Cambridge University Press, Cambridge (1997)
2. Ganter, B., Wille, R.: Formal Concept Analysis: Mathematical Foundations. Springer, New York (1999)
3. Goguen, J.A.: Data, schema, ontology and logic integration. Log. Jnl. IGPL. vol. 13, pp. 685–715. Oxford University Press (2005)
4. Goguen, J.: Information Integration in Institutions. Draft paper for the Jon Barwise memorial volume edited by Larry Moss (2006).
5. Goguen, J., Burstall, R.: Institutions: Abstract Model Theory for Specification and Programming. J. Assoc. Comp. Mach. 39, 95–146 (1992)
6. Kent, R.E.: Semantic Integration in the IFF. [[http://ftp.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-82/SI\\_paper\\_09.pdf](http://ftp.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-82/SI_paper_09.pdf)] In: Doan, A., Halevy, A., Noy, N., (eds.) Semantic Integration 2003. CEUR Workshop Proceedings, vol. 82, Sun SITE Central Europe (CEUR) (2003)
7. Kent, R.E.: Semantic Integration in the Information Flow Framework. In: Kalfoglou, Y., Schorlemmer, M., Sheth, A., Staab, S., Uschold, M. (eds.) Semantic Interoperability and Integration, Dagstuhl Seminar Proceedings, vol. 04391, Dagstuhl Research Online Publication Server (2005).
8. The Information Flow Framework (IFF), [<http://suo.ieee.org/IFF/>].